On Digital Communication Over a Discrete-Time Gaussian Channel with Noisy Feedback

By AARON D. WYNER

(Manuscript received April 21, 1969)

We consider the problem of transmission of digital data over a discrete-time Gaussian channel with the use of a Gaussian feedback channel. We are particularly interested in the case where the signal-to-noise ratio in the feedback channel is finite. By making use of simple extension of P. Elias' scheme for transmitting analog data over this channel with feedback, we show that it is possible at some transmission rates to increase the error-exponent (reliability) compared to the error-exponent found by C. E. Shannon for the one-way channel. In particular at transmission rate zero, we show that the error-exponent can be improved by a factor of $1 + [\hat{\rho}/(1 + \rho)]$, *where* ρ *and* $\hat{\rho}$ *are the forward and feedback signal-to-noise ratios respectively.*

I. INTRODUCTION

We consider the problem of transmission of digital data over a discrete-time Gaussian channel with the use of a Gaussian feedback channel. We are particularly interested in the case where tbe signal-to-noise ratio in the feedback channel is finite.

In Sections II and III we consider Elias' scheme and a simple extension for transmitting analog data over this channel with feedback.[1,2] In Section IV we apply this extended Elias scheme to the digital transmission problem. The main result is that for any rate $R < R^*$, a number less tban the channel capacity, it is possible to transmit digital data at a rate R with error probability

$$P_e = \exp\left[-E^* n_o + o(n_o)\right], \quad \text{as} \quad n_o \to \infty,$$

where n_o is the encoding-decoding delay, and $E^* > E_1$, tbe "one-way" exponent estimated by Shannon.[3] In particular, wben $R = 0$, $E_1 = \rho/4$ and $E^* = (\rho/4)[1 + \hat{\rho}(1 + \rho)^{-1}]$, where ρ and $\hat{\rho}$ are the forward and

feedback signal-to-noise ratios respectively. Finally, we suggest a modfication of this scheme which will probably permit extending R^* to capacity.

Stimulated by the work of Schalkwijk and Kailath, a great deal of research has been done on this problem (see for example Refs. 4–11). To the present author's knowledge, however, the result in this paper is the first to show that a noisy feedback channel can improve the error-exponent for digital communication on a band-limited channel. (References 4 and 8 treat the infinite band case.) Like the optimal coding schemes for the one-way channel, our scheme is not constructive. Let us remark here that this discrete-time channel is a model for the continuous-time Gaussian channel with a bandwidth constraint. (See Ref. 12 or 13.)

II. STATEMENT OF ELIAS' PROBLEM

We define a Gaussian channel as follows. The input is a real number x and the output is a number $y = x + z$, where the "noise" z is a Gaussian variate with mean zero and variance σ^2 and is independent of x. We assume here that the channel input x is a random variable, and require that the expectation $Ex^2 \leq P$, the "signal power".

To begin with, let us suppose that we wish to transmit the value of a random variable θ with the use of N transmissions over a Gaussian channel (with parameters P and σ^2). Assume also that a feedback Gaussian channel (with parameters $\hat{P}$ and $\hat{\sigma}^2$) is available which we may use $(N - 1)$ times alternating with the N forward uses. We assume nothing about the statistical nature of θ except that the expectation $E\theta^2 = \sigma_\theta^2$. Our goal is to obtain an unbiased estimate $\hat{\theta}$ of θ with minimum possible mean-squared error. Further, we restrict ourselves to linear processing of all data. We now state the problem and constraints precisely.

The forward and feedback channels are memoryless Gaussian channels with signal power P and $\hat{P}$ respectively and noise power σ^2 and $\hat{\sigma}^2$ respectively. Thus for the nth use of the forward channel the input is x_n and the output is $y_n = x_n + z_n$, where $Ex_n^2 = P$ and z_n is a Gaussian variate (independent of x_n) with mean zero and variance σ^2. For the nth use of the feedback channel the input is $\hat{x}_n$ and the output is $\hat{y}_n = \hat{x}_n + \hat{z}_n$, where $E\hat{x}_n^2 = \hat{P}$ and $\hat{z}_n$ is a Gaussian variate (independent of $\hat{x}_n$) with mean zero and variance $\hat{\sigma}^2$. We assume that the random variables $\{\theta, z_n, \hat{z}_n\}$ are independent. The condition requiring "linear processing" means the following. The input x_n to the forward channel (at the nth use) is given by

$$x_1 = a_1\theta, \qquad x_n = a_n\theta + \sum_{k=1}^{n-1} b_{nk}\hat{y}_k\,, \qquad n = 2, 3, \cdots, N. \tag{1}$$

The input to the feedback channel $\hat{x}_n$ (at the nth use) is given by

$$\hat{x}_n = \sum_{k=1}^{n} c_{nk}y_k\,, \qquad n = 1, 2, \cdots, N-1. \tag{2}$$

Finally, the receiver's estimate after N uses of the forward channel (and $N - 1$ uses of the feedback channel) is

$$\hat{\theta} = \sum_{n=1}^{N} d_n y_n\,. \tag{3}$$

We require that $\hat{\theta}$ be unbiased, that is, that given that $\theta = \theta_o$, the conditional expectation of $\hat{\theta}$ is

$$E(\hat{\theta} \mid \theta = \theta_o) = \theta_o\,. \tag{4}$$

The mean squared-error, which we wish to minimize is

$$\gamma^2 = E(\theta - \hat{\theta})^2. \tag{5}$$

Let γ^2_{OPT} be the minimum attainable value of γ^2 (over all choices of the coefficients a_n, b_{nk}, c_{nk}, d_n). It is easy to show that

(*i*) γ^2_{OPT} depends on P and σ^2 only through their ratio $\rho \triangleq P/\sigma^2$ (the forward "signal-to-noise" ratio), and on $\hat{P}$ and $\hat{\sigma}^2$ only through $\hat{\rho} \triangleq \hat{P}/\hat{\sigma}^2$.

(*ii*) for a given N, ρ, and $\hat{\rho}$, γ^2_{OPT} is proportional to σ^2_θ.

Thus we can write

$$\gamma^2_{\mathrm{OPT}} = \sigma^2_\theta \epsilon^2_{\mathrm{OPT}}(\rho, \hat{\rho}, N),$$

and our problem reduces to the determination of $\epsilon^2_{\mathrm{OPT}}(\rho, \hat{\rho}, N)$ (which can be thought of as a noise-to-signal ratio).

Let us observe that from the linearity assumptions (equation 1, 2, and 3) it follows that

$$\hat{\theta} = a\theta + \xi, \tag{6}$$

where a is a constant and ξ is a Gaussian variate independent of θ. From equation (4) it follows that $a = 1$ and $E\xi = 0$, and from equation (5) $E\xi^2 = \gamma^2$. Thus we can rewrite equation (6) as

$$\hat{\theta} = \theta + \xi, \tag{7}$$

where ξ is a Gaussian variate (independent of θ) with mean zero and variance γ^2. The important point here is that the entire process may

he thought of as reducing the N uses of the forward channel (and the $N - 1$ uses of the feedback channel) to a single one-way Gaussian channel with signal-to-noise ratio $(E\theta^2)/(E\xi^2) = \sigma_\theta^2/\gamma^2$.

III. ELIAS' RESULT

Elias solved our problem for the special case $N = 2$, where two uses of the forward channel and one of the feedback channel are permitted.[1,2] In his solution Elias admits the possibility that for the two uses of the forward channel, the signal-to-noise ratios are ρ_1 and ρ_2 respectively, where ρ_1 is not necessarily equal to ρ_2. His result is that the smallest attainable mean-squared error is given by

$$\gamma_E^2 = \sigma_\theta^2 \left[\rho_1 + \rho_2 + \frac{\rho_1 \rho_2 \hat{\rho}}{(1 + \rho_1)(1 + \rho_2) + \hat{\rho}} \right]^{-1}. \tag{8}$$

As discussed at the end of Section II, we can consider the entire process as a single one-way gaussian channel with signal-to-noise ratio $\sigma_\theta^2/\gamma_E^2$. We now turn to our problem, and note that we can obtain a (suboptimal) solution by applying Elias' technique recursively. For $N = 2$ we can, by setting $\rho_1 = \rho_2 = \rho$ in equation (8), obtain a signal-to-noise ratio $S_2 = \{2\rho + \rho^2\hat{\rho}/[(1 + \rho)^2 + \hat{\rho}]\}$. For $N = 3$ we can, by setting $\rho_1 = S_2$ and $\rho_2 = \rho$, obtain a signal-to-noise ratio S_3 given by

$$S_3 = \left[S_2 + \rho + \frac{\rho \hat{\rho} S_2}{(1 + S_2)(1 + \rho) + \hat{\rho}} \right],$$

and for arbitrary N we can obtain a signal-to-noise ratio S_N given by the recurrence

$$S_N = S_{N-1} + \rho + \frac{\rho \hat{\rho} S_{N-1}}{(1 + S_{N-1})(1 + \rho) + \hat{\rho}}, \tag{9a}$$

with initial condition

$$S_1 = \rho. \tag{9b}$$

Although equation (9) is difficult to solve explicitly we can obtain an approximate solution valid for large N. From equation (9a)

$$S_{N-1} + \rho \leqq S_N \leqq S_{N-1} + \rho + \frac{\hat{\rho}\rho}{(1 + \rho)}, \tag{10}$$

so that

$$\rho N \leqq S_N \leqq \left(\rho + \frac{\hat{\rho}\rho}{1 + \rho} \right) N. \tag{11}$$

We will show that as $N \to \infty$, S_N is asymptotic to the right member of inequality (11). Let us rewrite equation (9a)

$$S_{N+1} = S_N + \rho + \frac{\rho\hat{\rho}}{(1 + \rho)}\left[1 + \frac{1 + \rho + \hat{\rho}}{(1 + \rho)S_N}\right]^{-1}. \tag{12}$$

Let $S_N = [\rho + \hat{\rho}\rho/(1 + \rho)]N + \delta_N$, and expand the last term in equation (12) into a power series in $(1 + \rho + \hat{\rho})/[(1 + \rho)S_N]$. We then obtain, after cancelling terms,

$$\delta_{N+1} = \delta_N + \frac{\rho\hat{\rho}}{(1 + \rho)}\left[-\frac{(1 + \rho + \hat{\rho})}{(1 + \rho)S_N} + \frac{(1 + \rho + \hat{\rho})^2}{(1 + \rho)^2 S_N^2} + \cdots\right]. \tag{13}$$

From equation (11) we have that $S_N = O(N)$, so that equation (13) becomes

$$\delta_{N+1} - \delta_N = -O(1/N), \tag{14}$$

and therefore

$$\delta_N = -O(\log N). \tag{15}$$

Thus we conclude that

$$S_N = \left[\rho + \frac{\rho\hat{\rho}}{(1 + \rho)}\right]N - O(\log N). \tag{16}$$

An exact solution for S_N for various values of ρ, $\hat{\rho}$, and N is given in Table I. S_N^{-1} provides an upper bound to $\epsilon^2_{\mathrm{OPT}}$.

Elias also found a lower bound to $\epsilon^2_{\mathrm{OPT}}$,

$$\epsilon^2_{\mathrm{OPT}} \geqq 1/[\rho N + \hat{\rho}(N - 1)]. \tag{17}$$

This is the mean-squared error which results when the feedback channel is reversed and used in the forward direction, and we are allowed to use the forward channel N times and the feedback channel $(N - 1)$ times. Combining these results we have that

$$[(\rho + \hat{\rho})N]^{-1} \leqq \epsilon^2_{\mathrm{OPT}} \leqq \left[\left(\rho + \frac{\rho\hat{\rho}}{1 + \rho}\right)N - O(\log N)\right]^{-1}. \tag{18}$$

Let us remark here that the recurrence (9) can be solved exactly for the special case $\hat{\rho} = \infty$. In this case equation (9a) becomes

$$S_N = S_{N-1}(1 + \rho) + \rho, \tag{19}$$

and the solution is

$$S_N = (1 + \rho)^N - 1. \tag{20}$$

TABLE I—THE EXTENDED ELIAS SCHEME

FORWARD SNR $= \rho$, FEEDBACK SNR $= \hat{\rho}$

ASYMP. SNR $= \lim_{N\to\infty} \frac{S_N}{N} = \rho + \frac{\rho\hat{\rho}}{1+\rho}$, ASYMP. $E(0) = E^*(0)$

FORWARD SNR = 0.01 FEEDBACK SNR = 0.01
ASYMP. SNR = 0.010099 ASYMP. $E(0)$ = 2.52475E-03†

N	EQ. SNR $= S_N$	EQ. $E(0) = E_N(0)$	CAPACITY $= c_N$
1	0.01	0.0025	4.97516E-03
2	0.020001	2.50012E-03	4.95089E-03
3	3.00029E-02	2.50024E-03	4.92693E-03
4	4.00058E-02	2.50036E-03	4.90328E-03
5	5.00095E-02	2.50048E-03	4.87992E-03
6	6.00142E-02	2.50059E-03	4.85686E-03
7	7.00197E-02	2.50071E-03	4.83408E-03
8	8.00262E-02	2.50082E-03	4.81158E-03
9	9.00334E-02	2.50093E-03	4.78935E-03
10	0.100042	2.50104E-03	4.76740E-03

FORWARD SNR = 0.01 FEEDBACK SNR = 0.1
ASYMP. SNR = 1.09901E-02 ASYMP. $E(0)$ = 2.74752E-03

N	EQ. SNR $= S_N$	EQ. $E(0) = E_N(0)$	CAPACITY $= c_N$
1	0.01	0.0025	4.97516E-03
2	2.00089E-02	2.50112E-03	4.95284E-03
3	3.00266E-02	2.50222E-03	4.93078E-03
4	0.040053	2.50331E-03	4.90895E-03
5	5.00878E-02	2.50439E-03	4.88738E-03
6	6.01309E-02	2.50546E-03	4.86604E-03
7	7.01823E-02	2.50651E-03	4.84493E-03
8	8.02417E-02	2.50755E-03	4.82405E-03
9	9.03091E-02	2.50859E-03	4.80340E-03
10	0.100384	2.50961E-03	4.78297E-03
20	0.20154	2.51924E-03	4.59009E-03
30	0.303354	2.52795E-03	4.41569E-03
40	0.40574	2.53587E-03	4.25704E-03
50	0.508622	2.54311E-03	4.11197E-03
100	1.02871	2.57178E-03	3.53701E-03
150	1.55532	2.59219E-03	3.12725E-03
200	2.0861	2.60762E-03	2.81727E-03
250	2.61979	2.61979E-03	2.57283E-03
300	3.1556	2.62967E-03	2.37410E-03
350	3.69305	2.63789E-03	2.20869E-03
400	4.23179	2.64487E-03	2.06844E-03
450	4.77156	2.65087E-03	1.94771E-03
500	5.31219	2.65609E-03	1.84248E-03

† The notation "3E-5" means 3×10^{-5}.

TABLE I—(Continued)

FORWARD SNR = 0.01 FEEDBACK SNR = 1
ASYMP. SNR = 0.019901 ASYMP. $E(0)$ = 4.97525E-03

N	EQ. SNR = S_N	EQ. $E(0)$ = $E_N(0)$	CAPACITY = c_N
1	0.01	0.0025	4.97516E-03
2	2.00495E-02	2.50619E-03	4.96279E-03
3	3.01483E-02	2.51235E-03	4.95045E-03
4	0.040296	2.51850E-03	4.93816E-03
5	5.04925E-02	2.52463E-03	4.92591E-03
10	0.102197	2.55493E-03	4.86529E-03
15	0.155082	2.58470E-03	4.80572E-03
20	0.209115	2.61394E-03	4.74722E-03
40	0.436077	2.72548E-03	4.52394E-03
60	0.678846	2.82852E-03	4.31755E-03
80	0.935508	2.92346E-03	4.12731E-03
100	1.20434	3.01085E-03	3.95214E-03
300	4.31176	3.59313E-03	2.78320E-03
500	7.79234	3.89617E-03	2.17388E-03
700	11.4295	4.08196E-03	1.80005E-03
900	15.1502	4.20840E-03	1.54552E-03

FORWARD SNR = 1 FEEDBACK SNR = 0.1
ASYMP. SNR = 1.05 ASYMP. $E(0)$ = 0.2625

N	EQ. SNR = S_N	EQ. $E(0)$ = $E_N(0)$	CAPACITY = c_N
1	1	0.25	0.346574
2	2.02439	0.253049	0.276677
3	3.05731	0.254776	0.23342
4	4.09453	0.255908	0.203521
5	5.13433	0.256716	0.18139
6	6.17584	0.257327	0.164227
7	7.21857	0.257806	0.150457
8	8.26222	0.258194	0.139122
9	9.30658	0.258516	0.129599
10	10.3515	0.258788	0.121468

FORWARD SNR = 1 FEEDBACK SNR = 1
ASYMP. SNR = 1.5 ASYMP. $E(0)$ = 0.375

N	EQ. SNR = S_N	EQ. $E(0)$ = $E_N(0)$	CAPACITY = c_N
1	1	0.25	0.346574
2	2.2	0.275	0.290788
3	3.4973	0.291441	0.250579
4	4.84722	0.302951	0.220746
5	6.22905	0.311453	0.197811
6	7.63202	0.318001	0.179623
7	9.04989	0.32321	0.164826
8	10.4788	0.327462	0.152531
9	11.9162	0.331005	0.142138
10	13.3603	0.334007	0.133223

TABLE I—(Continued)

FORWARD SNR = 1   FEEDBACK SNR = 100
ASYMP. SNR = 51   ASYMP. $E(0)$ = 12.75

N	EQ. SNR = S_N	EQ. $E(0)$ = $E_N(0)$	CAPACITY = c_N
1	1	0.25	0.346574
2	2.96154	0.370192	0.344158
3	6.70566	0.558805	0.340326
4	13.5159	0.844743	0.334405
5	24.9907	1.24954	0.325774
6	42.434	1.76808	0.31427
7	66.142	2.36222	0.300486
8	95.3736	2.98042	0.285515
9	128.952	3.58201	0.270398
10	165.782	4.14455	0.255834
50	2073.46	10.3673	7.63746E-02
90	4079.34	11.3315	4.61885E-02
200	9646.14	12.0577	0.022936

FORWARD SNR = 100   FEEDBACK SNR = 1
ASYMP. SNR = 100.99   ASYMP. $E(0)$ = 25.2475

N	EQ. SNR = S_N	EQ. $E(0)$ = $E_N(0)$	CAPACITY = c_N
1	100	25	2.30756
2	200.98	25.1225	1.32704
3	301.965	25.1638	0.95227
4	402.952	25.1845	0.750162
5	503.94	25.197	0.622444
6	604.928	25.2053	0.533897
7	705.916	25.2113	0.468637
8	806.905	25.2158	0.418403
9	907.894	25.2193	0.378457
10	1008.88	25.2221	0.345879

Since the capacity, $\frac{1}{2}\log(1 + S_N)$, of the equivalent Gaussian channel (with signal-to-noise ratio S_N), cannot exceed N times the capacity, $\frac{1}{2}\log(1 + \rho)$, of a single channel (with signal-to-noise ratio ρ), S_N as given by equation (20) is in fact optimal. Thus

$$\epsilon^2_{\text{OPT}}(\rho, \infty, N) = [(1 + \rho)^N - 1]^{-1}, \tag{21}$$

which is an exponential in N.

IV. APPLICATION TO DIGITAL COMMUNICATION

4.1 *Schalkwijk–Kailath Technique*

Suppose we wish to transmit one of M equally likely messages over a Gaussian forward channel with signal-to-noise ratio ρ with the aid of

TABLE I—(Continued)

FORWARD SNR = 100 ASYMP. SNR = 199.01		FEEDBACK SNR = 100 ASYMP. $E(0)$ = 49.7525	
N	EQ. SNR = S_N	EQ. $E(0) = E_N(0)$	CAPACITY = c_N
1	100	25	2.30756
2	297.078	37.1347	1.42434
3	495.429	41.2858	1.03457
4	694.043	43.3777	0.817997
5	892.77	44.6385	0.679545
6	1091.56	45.4816	0.583023
7	1290.39	46.0853	0.511677
8	1489.25	46.539	0.456669
9	1688.12	46.8923	0.412887
10	1887.02	47.1754	0.377164

FORWARD SNR = 100 ASYMP. SNR = 1090.1		FEEDBACK SNR = 1000 ASYMP. $E(0)$ = 272.525	
N	EQ. SNR = S_N	EQ. $E(0) = E_N(0)$	CAPACITY = c_N
1	100	25	2.30756
2	1092.78	136.597	1.74935
3	2173.1	181.091	1.28073
4	3258.25	203.641	1.01116
5	4345.00	217.253	0.837702
10	9786.78	244.669	0.459444
15	15232.7	253.878	0.321.042
20	20680.1	258.501	0.248424
40	42474.8	265.467	0.133209
60	64272.6	267.803	9.22575E-02
80	86071.7	268.974	7.10184E-02
100	107871.	269.679	5.79435E-02
120	129672.	270.149	4.90532E-02
140	151472.	270.486	4.26006E-02
160	173273.	270.738	3.76957E-02
180	195073.	270.935	3.38365E-02
200	216874.	271.093	3.07177E-02

a Gaussian feedback channel with signal-to-noise ratio $\hat{\rho}$, using the forward channel N times and the feedback channel $N - 1$ times. Following Schalkwijk and Kailath,[10] we assign to message i ($i = 1, 2, \cdots, M$) the number $\theta = \theta_i = i - (M + 1)/2$. Thus the M messages are equally spaced on the interval $[-(M - 1)/2, (M - 1)/2]$ at distance 1 apart. We can now apply the results of Sections III and IV to transmit θ. The expectation $E\theta^2 = \sigma_\theta^2$ is

$$\sigma_\theta^2 = (M + 1)(M - 1)/12, \quad M = 1, 2, 3 \cdots . \tag{22}$$

When message i is transmitted, the output of the system is $\hat{\theta} = \theta_i + \xi$, where ξ is a zero mean Gaussian random variable with variance γ^2. We select as the decoder output, that $j(1 \leqq j \leqq M)$ which minimizes $|\hat{\theta} - \theta_j|$, so that we make an error only when $|\xi| \geqq \frac{1}{2}$. This event has probability

$$P_e = 2\Phi(-1/2\gamma), \tag{23}$$

where $\Phi(x) = 1/(2\pi)^{\frac{1}{2}} \int_{-\infty}^{x} \exp(-x^2/2)\, dx$ is the cumulative error function. Thus the smallest error probability attainable using this scheme (with parameters N, ρ, $\hat{\rho}$) is

$$P_{e,\,\mathrm{OPT}} = 2\Phi\left[-\frac{1}{2\epsilon_{\mathrm{OPT}}(\rho,\, \hat{\rho},\, N)\sigma_\theta}\right] \tag{24}$$

where σ_θ is given by equation (22) and ϵ_{OPT} in Section II. The bounds on $\epsilon_{\mathrm{OPT}}^2$ in Section III immediately yield bounds on $P_{e,OPT}$.

Let us assume that every T seconds, a digital message source emits one of $M = e^{RT}$ equally likely messages (R is the message "rate"). Further assume that $N = \alpha T$ (for example, if the "physical" channel has bandwidth W cps, then $\alpha = 2W$). Consider two cases: $\hat{\rho} = \infty$, $\hat{\rho} < \infty$.

(i) When $\hat{\rho} = \infty$, it follows immediately from equation (21) and (22) that as $T \to \infty$

$$\frac{1}{2\epsilon_{\mathrm{OPT}}(\rho,\, \infty,\, N)\sigma_\theta} \sim \sqrt{3}\, \frac{(1 + \rho)^{\alpha T}}{e^{RT}} = \sqrt{3}\, e^{(C-R)T}, \tag{25}$$

where $C = (\alpha/2) \log (1 + \rho)$ is the channel capacity in nats per second. Thus, provided $R < C$, as $T \to \infty$ the argument of Φ in equation (24) becomes infinite and $P_{e,\mathrm{OPT}} \to 0$. In fact, (since $\Phi(x) \sim (2\pi x^2)^{-\frac{1}{2}} \exp(-x^2/2)$, as $x \to \infty$)

$$P_{e,\,\mathrm{OPT}} = \exp\left[e^{2(C-R)T+o(T)}\right], \quad \text{as} \quad T \to \infty, \tag{26}$$

a double exponential decay. This is the celebrated result of Schalkwijk and Kailath.[10,11]

(ii) If we try to apply the same scheme when the feedback signal-to-noise $\hat{\rho} < \infty$, then from equation (18) $(2\epsilon_{\mathrm{OPT}}\sigma_\theta)^{-1} \to 0$ as $T \to \infty$. Thus it is not possible using this scheme to obtain vanishingly small error probability as $T \to \infty$ with fixed signal-to-noise ratios in the forward and feedback channel. This is so no matter how large $\hat{\rho}$ may be, provided it is finite. For finite T however, equations (18) and (24) yield useful estimates of attainable error probabilities.

4.2 *Improving the One-Way Error Exponent*

Suppose that, as in Section 4.1, we wish to transmit one of $M = e^{RT}$ equally likely messages in T seconds. Suppose that we use only a forward Gaussian channel (with signal-to-noise ratio ρ) $n_o = \alpha T$ times. Then it is well known that one can attain an error probability

$$P_e = \exp\left[-E_1\left(\frac{R}{\alpha}, \rho\right)\alpha T + o(T)\right], \quad \text{as} \quad T \to \infty, \tag{27}$$

where $E_1(R/\alpha, \rho) > 0$, if $R < \alpha/2 \log(1 + \rho) = C$, the channel capacity. As indicated, the quantity $E_1(R/\alpha, \rho)$ depends on R and α only through their ratio. Although E_1 is not known exactly, estimates are given in Ref. 3.† In particular, $E_1(0, \rho) = \rho/4$ and $E_1(C/\alpha, \rho) = 0$.

Now suppose we have a Gaussian feedback channel available with signal-to-noise ratio $\hat{\rho}$. Let us divide the n_o forward channel uses into $\nu = n_o/N$ groups of N forward channel uses. In each of these groups we use the extended Elias scheme, (of Sections III and IV, with N uses of the forward channel and $N - 1$ uses of the feedback channel) to generate an equivalent forward Gaussian channel with signal-to-noise ratio S_N given by the recurrence (9). We then use a one-way coding scheme with $\nu = n_o/N = (\alpha/N)T$ uses of the equivalent forward channel. With N held fixed as $T \to \infty$, we can attain an error probability as in equation (27) with α replaced by (α/N) and ρ replaced by S_N—namely,

$$P_e = \exp\left[-E_1\left(\frac{RN}{\alpha}, S_N\right)\frac{\alpha T}{N} + o(T)\right]. \tag{28}$$

Thus the new error-exponent is

$$E_N(R, \rho, \hat{\rho}) = \frac{1}{N} E_1\left(\frac{RN}{\alpha}, S_N\right). \tag{29}$$

Since N is arbitrary, we can state our result:

Theorem: Given a forward and feedback Gaussian channels which can each process α inputs (independently) per second, with signal-to-noise ratio ρ and $\hat{\rho}$ respectively. Then it is possible to transmit digital data at a rate R nats per second with error probability

$$P_e = \exp\left[-E^*\alpha T + o(T)\right], \qquad \textit{as} \qquad T \to \infty, \tag{30a}$$

† The conventional power constraint for a one-way channel is that the time average of the square of the inputs must not exceed P. The power constraint used here is that the *statistical* expectation of the square of *each* input not exceed P. Nither of these constraints imply the other. However, it is not hard to show that the estimates of E_1 (in Ref. 3) are valid for both constraints.

where

$$E^* = E^*\left(\frac{R}{\alpha}, \rho, \hat{\rho}\right) = \sup_{1 \leq N < \infty} E_N = \sup_{1 \leq N < \infty} \frac{1}{N} E_1\left(\frac{RN}{\alpha}, S_N\right), \tag{30b}$$

S_N is the solution to the recurrence (9), and E_1 is the reliability (error-exponent) for the one-way Gaussian channel as in equation (27), and T is the encoding-decoding delay.

Remarks:

(*i*) Since $S_1 = \rho$ and $E_1(R/\alpha, \rho) > 0$ for $R < \alpha/2 \log (1 + \rho) = C$, then $E^*(R/\alpha, \rho, \hat{\rho}) > 0$ for $R < C$.

(*ii*) Since $E_1(0, \rho) = \rho/4$,

$$E_N(0, \rho, \hat{\rho}) = \frac{1}{N} E_1\left(\frac{RN}{\alpha}, S_N\right)\Big|_{R=0} = \frac{S_N}{4N},$$

so that from equation (16),

$$E^*(0, \rho, \hat{\rho}) \geqq \frac{S_N}{4N} \rightarrow \frac{\rho}{4}\left(1 + \frac{\hat{\rho}}{1 + \rho}\right), \quad \text{as} \quad N \rightarrow \infty.$$

In fact, since S_N/N can be shown to be non-decreasing, $E^*(0, \rho, \hat{\rho})$ is in fact equal to this quantity. Thus the use of the feedback channel represents an improvement of a factor of $[1 + \hat{\rho}/(1 + \rho)]$ in the error-exponent at zero rate.

(*iii*) We can get a rough idea of the behavior of $E^*(R/\alpha, \rho, \hat{\rho})$ as follows. Let $r = R/\alpha$ be the rate in nats per channel use. Let us *crudely* approximate the one-way exponent $E_1(r, \rho)$ as r varies from 0 to $c = C/\alpha$ (the capacity in nats per channel use) by a straight line connecting $(r = 0, E_1 = \rho/4)$ and $(r = c, E_1 = 0)$. See Fig. 1.
Then E_2 has $r = 0$ intercept at

$$E_2(0, \rho, \hat{\rho}) = \frac{S_2}{2 \cdot 4} = \frac{\rho}{4} + \frac{\hat{\rho}\rho^2}{8[(1 + \rho)^2 + \hat{\rho}]} > \frac{\rho}{4},$$

and $E_2(r, \rho, \hat{\rho}) = 0$ at $r = c_2 \triangleq (1/2)(1/2) \log (1 + S_2)$. Similarly, E_N has $r = 0$ intercept at

$$E_N(0, \rho, \hat{\rho}) = \frac{SN}{4N}, \frac{SN}{4(N - 1)},$$

and $E_N(r, \rho, \hat{\rho}) = 0$ at $r = c_N \triangleq (1/2N) \log (1 + S_N)$. From Fig. 1, we see that for each value of $r > 0$, there is a value of $N (1 \leqq N < \infty)$ which maximizes $E_N(r, \rho, \hat{\rho})$ to achieve $E^*(r, \rho, \hat{\rho})$. Values of $E_N(0, \rho, \hat{\rho})$ and c_N are tabulated for various values of ρ, $\hat{\rho}$, and N in Table I.

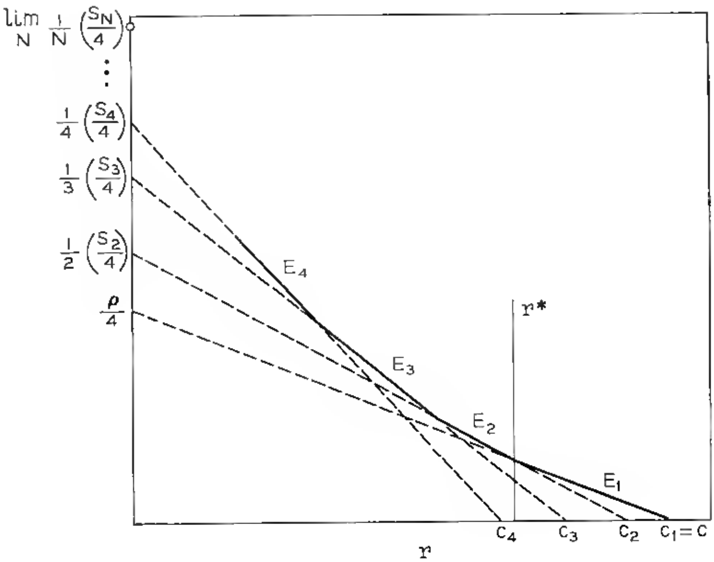

Fig. 1 — $E^*(r, \rho, \hat{\rho})$ vs. r (an approximation).

(*iv*) We see from Fig. 1, that the feedback scheme offers no improvement over the one-way scheme (that is, $E_N < E_1$, for $N > 2$) for $r^* \leqq r < c$ where r^* is the solution of $E_2 = E_1$, that is,

$$\tfrac{1}{2}E_1(2r^*, S_2) = E_1(r^*, \rho).$$

However, the rate $r^* \to c$ as $\hat{\rho} \to \infty$.

Actually, it is probably possible to improve on our results substantially and in particular bring about an increase in the error-exponent for all $r < c$. Let $\{N_1, N_2, \cdots, N_k\}$ be a set of positive integers (not necessarily equal). Then divide the $n_o = \alpha T$ forward channel uses into $\nu = n_o/(N_1 + N_2 + \cdots N_k)$ uses of an equivalent channel which is the parallel combination of k Gaussian channels with signal-to-noise ratios $S_{N_1}, S_{N_2}, \cdots, S_{N_k}$. These k Gaussian channels are generated by $N_1, N_2, \cdots, N_k$ iterations, respectively, of the Elias scheme. One must then compute the error-exponent for a parallel combination of channels to obtain a new improved exponent.[14] We leave this task as an open problem.

(*v*) Let us finally remark that although the expectation of the cbannel input power x^2 is constrained, the quantity x^2 is in fact a random variable distributed on the interval $[0, \infty)$. This is in contrast to the one-way

schemes where the channel input is bounded. This point is discussed in Ref. 13.

REFERENCES

1. Elias, P., "Channel capacity without coding," *Lectures on Communication System Theory*, New York: McGraw Hill. 1961, pp. 363–366.
2. Elias, P., "Networks of Gaussian channels with application to feedback systems," IEEE Trans. Inform. Theory, *IT-13*, No. 7 (July 1967), pp. 493–501.
3. Shannon, C. E., "Probability of error for optimal codes in a Gaussian channel," B.S.T.J., *38*, No. 3 (May 1959), pp. 611–656.
4. Butman, J., "Phase in coherent feedback communication," paper presented at 1969 Int. Symp. on Inform. Theory, Ellenville, New York.
5. Butman, S., "Optimum linear coding for additive noise systems using information feedback." Technical Rep. No. 1, Communication Theory Laboratory, California Institute of Technology, Pasadena, California, May 1967.
6. Kashyap, R. L., "Feedback coding schemes for an additive noise channel with a noisy feedback link," IEEE Trans. Inform. Theory, *IT-14*, No. 5 (May 1968), pp. 471–486.
7. Kramer, A. J., "Improving communication reliability by use of an intermittent feedback link," IEEE Trans. on Information Theory, *IT-15*, No. 1 (January 1969), pp. 52–60.
8. Lavenberg, S., "Feedback communication using orthogonal signals," IEEE Trans. on Information Theory, *IT-15*, No. 4 (July 1969), pp. 478–483.
9. Omura, J. K., "Optimum linear transmission of analog data for channels with feedback," IEEE Trans. Inform. Theory, *IT-14*, No. 1 (January 1968), pp. 38–43.
10. Schalkwijk, J. P. M., and Kailath, T., "A coding scheme for additive channels with feedback, part I," IEEE Trans. Inform. Theory, *IT-12*, No. 4 (April 1966), pp. 172–182.
11. Schalkwijk, J. P. M., "A coding scheme for additive channels with feedback, part II," IEEE Trans. Inform. Theory, *IT-12*, No. 4 (April 1966), pp. 183–189.
12. Wozencraft, J. M., and Jacobs, I. M., *Principles of Communication Engineering*, New York: Wiley, 1965, pp. 294–297.
13. Wyner, A. D., "On the Schalkwijk–Kailath coding scheme with a peak energy constraint," IEEE Trans. Inform. Theory, *IT-14*, No. 1 (January 1968), pp. 71–78.
14. Gallager, R., *Information Theory and Reliable Communication*, New York: Wiley, 1968, pp. 149–150.